6210-05 May 1994

Medical Staff Risk Pool Policies: Stability and Simulation

W. Pete Welch
Mark E. Miller

The Urban Institute
2100 M Street, N.W.
Washington, DC 20037

Support for this research was provided by the Health Care Financing Administration to The Urban Institute through Cooperative Agreement No. 18-C-90038. This Cooperative Agreement supported a number of research projects related to inhospital physician expenditures. The total budget for all these projects was $646,002 and represents the sole source of funding.

The authors would like to acknowledge Ellen Englert for research assistance, Paula Beasley and her staff at Social and Scientific Systems for programming, and John Marcotte for statistical advice. Any opinions expressed are those of the authors and do not necessarily reflect the opinions of the Health Care Financing Administration, The Urban Institute, or its sponsors.

TABLE OF CONTENTS

Abstract ii

I. Introduction 1
A. Motivations 1
B. Databases 2

II. Stability 3
A. Methods 3
B. Results 4
Regression Results 4
Distribution of First Differences 4
Random Fluctuation at the Admission Level 5
Correlations Based on Three Years of Data 6
C. Conclusion 6

III. Simulation of Illustrative Medical Staff Policies 7
A. Methods 7
Shrinkage Estimator 7
Regression Specification 10
Database 11
Regression Results 11
Shrinkage Estimates vs. Estimates Based on 100-Percent Data 12
B. Illustrative Simulations 13
Per-Case Prospective Payment 13
High-Cost Medical Staff Policy 14

References 16

Tables 17

Appendix A: Levels Vs. Growth Rates A-1
Appendix B: Random Fluctuation at the Admission Level B-1
Appendix C: Aggregation of Data on Volume per Admission C-1
Appendix D: Growth in Volume per Admission and the Casemix Index D-1

I. INTRODUCTION

This project investigates the effects of policies that might place medical staffs at risk for the volume and intensity of physician services provided in the hospital. Our measure of services is casemix-adjusted Relative Value Units (RVUs) per admission, which will be referred to as volume per admission.[1] This paper has two purposes: First, it investigates the year-to-year (or temporal) stability of volume per admission at the hospital level, and second, it simulates the impact of three medical staff policies.

In Welch and Miller (1992), we discussed three policy strategies: a Volume Performance Standard (VPS) growth standard for each medical staff, a per-case prospective payment to each staff (or per-admission budget), and a high-cost medical staff policy. Efforts to control hospital facility costs serve as precedents for these three medical staff strategies, all of which are aimed at controlling physician services per admission. The VPS policy would encompass all medical staffs, and fees would be adjusted upward or downward according to the growth in inpatient physician services per admission (the precedent here is the Tax Equity and Fiscal Responsibility Act of 1982 [TEFRA]). The second policy, also covering all medical staffs, would establish a per-admission budget for each medical staff (the precedent here is Medicare's Prospective Payment System [PPS]). This policy would adjust a medical staff's fees according to the difference between its average volume of services and the national average. The third policy would be targeted at medical staffs whose billings are substantially above the national average. Fees in these "high-cost" medical staffs would be adjusted downward (the precedent is Section 223 of the 1972 Medicare amendments).

The high-cost medical staff policy and the per-case prospective payment policy pertain to the level of physician services per admission, whereas the VPS policy pertains to the growth in physician services per admission. This paper concentrates on the two level policies, because of greater policy interest and because analysis of a growth policy requires three or four consecutive years of data, which we lack.[2]

A. Motivations

This paper is motivated, in part, by the following concern: If a hospital's volume per admission varies widely from year to year for little apparent reason, a medical staff's penalty or reward could be largely random and unrelated to its effort to control volume.[3] Great instability would fundamentally undermine a policy, whereas moderate instability might be mitigated, for example, by excluding small hospitals or by holding small hospitals to a higher limit than other hospitals. Section II finds that instability is not necessarily a serious problem for the two level policies

[1] In earlier work, we used the term "intensity."

[2] Appendix A argues that volume and intensity growth rates can be unstable when volume and intensity levels are quite stable. Hence, instability is a more serious problem for a staff-level VPS growth standard than for a per-admission payment or high-cost medical staff policy.

[3] Generally, the terms "hospital" and "medical staff" can be used interchangeably, because a hospital cannot function without a medical staff (i.e., its physicians) and a medical staff cannot function without a hospital. It is only when a hospital's physicians act corporately (e.g., reviewing practice patterns) that we use the term "medical staff."

(prospective payment and high-cost medical policy), but Appendix A shows that stability is more of a problem for a medical staff growth policy.

This paper is also motivated by the need to describe the impact of a medical staff policy. Because many policy proposals are motivated by the need for budget savings, an important aspect of a proposal is its budget savings. In addition, there are typically concerns regarding the distributional impacts of a policy. For instance, whenever a PPS policy option is proposed, policymakers wish to know which hospital categories would be helped or hurt. PPS impacts are often analyzed in terms of such hospital categories as number of beds, teaching status, and region of the country. This paper performs similar analyzes for most of these categories.

Estimating the budgetary impact is more difficult for a high-cost medical staff policy than for a per-case prospective payment policy. This is because the high-cost medical staff yields budget savings, in part, by lowering payment to medical staffs whose volume per admission exceeds the national mean. The greater the variance across hospitals (cross-sectionally, not temporally), the greater the savings from a policy that reduces payment to medical staffs that are, say, above 110 percent of the national mean. Whereas the savings from a per-case prospective payment policy depends mainly on the mean, the savings from a high-cost medical staff policy also depends on the cross-sectional variance (as well as the temporal variance).

B. Databases

As in the past, we linked physician claims to hospital stays by beneficiary ID and dates of service. (See Miller and Welch (1993) for details.) Data in this paper pertain to services delivered prior to the implementation of the Medicare Fee Schedule (MFS). Because RVUs are a direct measure of volume and intensity, we attached 1992 RVUs to each physician claim. For each 1989 Diagnosis Related Group (DRG), the RVUs per admission were computed with 1989 physician claims data. These national means were used to adjust each hospital's RVUs per admission for its casemix. For some analyses, outlier cases were truncated at 2.5 times the mean of the DRG to mitigate the incentive for a medical staff to discourage the admission of patients who are likely to be very expensive to treat.

The ideal database for these analyses would have been 100 percent of the physician bills in the nation in several years (under the MFS).[4] However, construction of such a database would entail much computer time. Instead, we have constructed two databases: a 5-percent national sample of beneficiaries and a 100-percent file for 7 states.[5] Each file has its own disadvantages.

---

[4]Data under the MFS has only recently become available.

[5]The 100-percent files for 1988 and 1989 were derived from files constructed by the Center for Health Economics Research (CHER). The source files had 100 percent of physician claims from 12 states (see CHER [1991] for details). To avoid excessive use of HCFA's computing resources, we limited our analyses to 7 states: Alabama, Arizona, Kansas, Oklahoma, Pennsylvania, Washington, and the Missouri component of the Kansas City Metropolitan Statistical Area. These states were selected in order to obtain diversity in terms of region of the country and urban-rural location. These states had about 13 percent of the U.S. population.

The 100-percent file for 1991 was derived from the National Claims History File. This is a 100-percent beneficiary file, whereas the CHER files are 100-percent claims files.

The disadvantage of the national file is that it is only a 5-percent sample. The estimator of volume per admission calculated at the hospital level has more variance when based on a 5-percent file than a 100-percent file, because the impact of randomness at the admission level decreases as the sample size increases.

The disadvantage of the 100-percent file is that it may exclude parts of the country that are sources of variance (e.g., the South Atlantic Census division). In particular, the 100-percent file excludes Florida, although south Florida has some of the highest volumes per admission in the country (Welch et al. 1993).

Given these considerations, we created 100-percent files for 7 states for three years: 1988, 1989, and 1991. We also created 5-percent files for all states for two years: 1989 and 1991. Section II uses the 100-percent files to analyze stability of volume per admission.

To analyze the impact of a policy on hospitals in all 50 states, we used the 5-percent files. Section III develops a methodology (shrinkage estimator) for calculating estimates of volume per admission using the 5-percent file. The 100-percent files for 1989 and 1991 are used to check on the reasonableness of this methodology. We conclude that, properly adjusted, the 5-percent files can be used to simulate policy impacts.

II. STABILITY

A. Methods

The result of our database construction was a hospital-level file with volume per admission in 1988, 1989, and 1991. For 733 hospitals, we had data for both 1988 and 1989. Most of our analyses are limited to 1988 and 1989, but at the end of this section we calculate correlations across the three years of data. For 714 hospitals, we had data for all three years. The latter database had the following admission counts and (unweighted by admissions) volumes per admission (i.e., casemix-adjusted RVUs per admission):

1988	1,218,194	29.8
1989	1,190,276	30.5
1991	1,290,738	34.4

We normalized volume per admission to 100 separately by year, so that a 1988 volume per admission of 100 represents the grand mean (weighted by admissions) for 1988 and a 1989 volume per admission of 100 represents the grand mean for 1989. This normalization factored out the 2.5 percent average growth between 1988 and 1989. Although average growth in volume could be the focus of a policy, our concern here is for changes relative to the grand mean, which represent the volume. For instance, suppose a hospital's volume per admission was constant over a three-year period during which the national mean (or grand mean) grew 10 percent per year. The hospital could have gone from 15 percent above the grand mean in year 1 to 8 percent below in year 3 even though in absolute terms its volume was constant. Note that our stability analyses are not weighted by admissions, because the relevant unit of observation from a policy perspective is the hospital.

B. Results

We regressed volume per admission in 1989 on volume per admission in 1988. In the extreme case of no change from one year to the next, the R-square would be 1.0 and the slope coefficient would be 1.0. If the R-square was considerably less than 1.0 but the slope was 1.0, the expected value of 1989 volume per admission (relative to the mean) would equal volume per admission (relative to the mean) in the previous year but there would be substantial variation around that mean (i.e., instability). Conversely, if the R-square was close to 1.0 but the slope was substantially less than 1.0, we could predict each hospital's 1989 volume per admission with precision based on its 1988 volume per admission. But volume per admission in the two years would be different, with volume per admission in the second year typically being closer to its year-specific grand mean.

Four variants of volume per admission were used in the regressions: stay only vs. stay plus windows, and truncated for outliers vs. untruncated (see Miller and Welch [1993]). The four variants allowed us to examine stability according to how two policy design decisions might be made. One design decision is whether medical staffs would be placed at risk only for services delivered during the stay or at risk also for services delivered immediately prior to admission or subsequent to discharge. The second design decision is whether a medical staff would be held responsible for all the physician services for a specific admission without limit or whether the staff would not be responsible for services beyond a per-admission threshold (i.e., an outlier policy).

Regression results

As shown in Table 1, the results vary little across definitions of volume per admission. Under all four definitions, the R-square is .94 (i.e., the correlation coefficient is .97). The slope coefficient is about .97, depending on the definition of volume per admission.

The slope coefficients are significantly different from 1.0, but the difference is quite small. This difference represents regression toward the mean: Whenever a hospital's volume per admission differs from the year-specific grand mean, part of that difference may be due to transitory factors, including random fluctuation at the admission level. By definition, transitory factors do not manifest themselves as perceptibly in the following year. Hence, whenever a hospital's volume per admission is, say, above the mean in the first year, it tends to decrease in the second year.

Distribution of first differences

Another useful measure of stability involves the change in volume per admission between 1988 and 1989 for each hospital. (Recall that volume per admission in each year was normalized. Only then was the first difference calculated.) Table 1 demonstrated that the choice of volume per admission variable (e.g., stay only or stay plus windows) does not affect the results. We focus on volume per admission for the stay, untruncated for outliers, as the most intuitive variant. (Because truncation lowers variances, the differences in this section would be slightly lower if they pertained to truncated values.)

As shown in Table 2, 10 percent of hospitals (or 73 hospitals) had volume per admission decrease by at least 7.2 percentage points, and 10 percent had volume per admission increase by

at least 6.1 percentage points. Put differently, about 80 percent of hospitals experienced changes of less than 7 percentage points of the national mean. The first difference of volume per admission had a standard deviation of 6.3 percentage points. This summary measure indicates that, if the distribution was normal, about two-thirds of the hospitals had a volume change (in absolute value) of less than 6.3 percentage points.

One policy design option is to exclude hospitals with fewer than 50 beds, in part because these hospitals may have few admissions. Table 2 also shows the distribution of first differences for hospitals with at least 50 beds (74 percent of the hospitals in our data). The standard deviation is 6.0 percentage points, in contrast to 6.3 percentage points for all hospitals and 7.0 percentage points for hospitals with 50 beds or fewer (not shown). Of these medium and large hospitals, 80 percent experienced changes of less than 7 percentage points.

Another useful subsample is staffs designated as high-cost based on 1988 volume per admission, because only these staffs would be affected by a high-cost medical staff policy. Although they tend to have many admissions per hospital, these hospitals also have a larger standard deviation than all hospitals: 6.8 vs. 6.3 percentage points. However, given regression toward the mean, this result is not surprising. High-cost staffs are, by definition, above the national mean. Their volume per admission is more likely to decrease than increase, as shown by the distribution.

Random fluctuation at the admission level

Variation in volume per admission can represent either random fluctuation at the admission level or systematic shifts at the hospital level. Within a hospital, some admissions have more casemix-adjusted RVUs than others. This random fluctuation at the admission level could result from one patient being more severely ill than another or from physicians having different styles of practice. Regardless of its cause, admission-level randomness is one reason why a hospital's measured volume per admission may vary between years.

How important is this source of hospital-level change? For each hospital in each year, we calculated the standard error of the estimate. We then calculated the t-value for the difference between normalized volume per admission in 1988 and 1989. (Recall that the average increase was factored out through normalization.)

Volume per admission was significantly different ($p<.05$) between the two years for 36 percent of the 733 hospitals. By way of contrast, this figure would be only 5 percent if a hospital's difference in volume per admission between years reflected only random fluctuation. Broken down by the absolute value of the difference in volume per admission between years, the results are as follows:

Absolute difference between 1988 and 1989 (percentage points)	Percent of hospitals for which the difference was significant
0 - 2	2%
2 - 4	21
4 - 6	60
6 - 8	84
>8	89

About 23 percent of hospitals experienced changes in volume per admission of at least 6 percentage points (Table 2). Eighty-seven percent of these differences are significant and hence cannot be attributed to fluctuation at the admission level. Hence, when a staff experiences a large change in volume per admission from one year to the next, it presumably is due to some structural or organizational change.

What causes a hospital's volume per admission to increase or decrease relative to the grand mean? It could be physicians joining or leaving the staff, the addition of an MRI scanner, or admission-level fluctuation. Naturally, this database does not enable us to consider these causes directly. However, for 36 percent of the hospitals, it does enable us to reject the hypothesis that admission-level fluctuation alone caused the change. For 87 percent of the hospitals experiencing a change of at least 6 percentage points, we can also reject the hypothesis. In those cases, there has apparently been some systematic change.

The standard error of the estimate (of the first difference), used in computing the t-value, is by itself a useful measure. The standard error for the typical hospital is .97 RVUs per admission, which is 2.7 percentage points of the 1988 average volume per admission. (See Appendix B.)

Suppose that the variation in volume per admission was solely the result of random fluctuation; that is. there were no systematic shifts in volume per admission over time. The standard deviation of the difference would be 2.7 percentage points, which is two-fifths of the actual standard deviation of 6.3 percentage points (Table 2). Hence, about 18 percent ($[2.7/6.3]^2$) of the within-hospital between-year variance in volume per admission can be attributed to admission-level fluctuation. The remaining variance is at the hospital level.

Correlations based on three years of data

As noted, when 1991 data are combined with 1988 and 1989 data, we have a three-year data series for 714 hospitals. Using these observations, we calculated the following pairwise correlations:

		untruncated	truncated
1-year	1988 and 1989	.973	.972
2-year	1989 and 1991	.916	.945
3-year	1988 and 1991	.909	.932

As one would expect, the correlation falls as the interval increases. However, even with a three-year difference, the correlation is still .909.

C. Conclusion

If a medical staff's volume per admission varied widely from year to year for little apparent reason, the staff might be penalized or rewarded without regard for its effort to control volume, raising issues of equity. Hence, we have investigated the temporal stability of volume per admission, which is relevant to the two level policies--a per-case prospective payment policy and a high-cost medical staff policy.

We conclude that (untruncated) volume per admission at the hospital level is not seriously unstable. This conclusion is based on several findings:

- Volume per admission in consecutive years is highly correlated (r=.97).
- The slope coefficient is slightly (but significantly) different from 1.0, indicating a little regression toward the mean.
- Year-to-year change in volume per admission exceeds 7 percentage points for one fifth of hospitals.
- Year-to-year change in volume per admission is only slightly greater for hospitals with fewer than 50 beds than those with more than 50 beds.

Although instability cannot be dismissed as a problem, these findings (particularly the correlation) are reassuring.

We also conclude that much of any year-to-year change appears systematic; that is, it is not the result of random fluctuation at the admission level. This conclusion is based on two findings:

- Eighty-seven percent of the time when the (absolute) change exceeds 6 percentage points, the change cannot be attributed solely to random fluctuation at the admission level.
- A minority (18 percent) of hospital-level variation is attributed to admission-level variance (see Appendix B).

To the extent that medical staffs would not be penalized by "luck" at the admission level, holding them responsible for their volume per admission could be justified despite some year-to-year variation.

III. SIMULATION OF ILLUSTRATIVE MEDICAL STAFF POLICIES

This section simulates the impact on medical staffs of two illustrative medical staff policies.

A. Methods

Shrinkage Estimator

We wish to estimate the volume per admission--casemix-adjusted RVUs per admission--for each hospital in, say, 1989. However, we have only a 5-percent sample of admissions for hospitals nationally. Hence, a given hospital may have a limited number of admissions (about 90 admissions, on average) upon which to base an estimate. The statistical problem is that an estimator based on the 5-percent sample will overestimate the variance in volume per admission across hospitals.

Our solution to this problem involves a shrinkage estimator (Efron and Morris, 1977), also labeled Stein or empirical Bayes estimator. Duan, Manning, and Newhouse (1989) and Miller et al. (1993) are examples of applications of this estimator to health care issues. In essence, a shrinkage estimator combines two estimators, each with their limitations. In the traditional example, a statistician has data on the batting averages for major league players for the first, say.

50 times at bat in a season. Although one could predict a player's season average from these data, the estimator would suffer from variance due to the small sample. Alternatively, one could predict a player's season average based on his average in the previous season. This estimator, however, would ignore any changes in batting ability from one year to the next. A third alternative is the shrinkage estimator, which is a weighted average of the two estimators. The weight for the conceptually most direct estimator (this season's average) decreases as its standard error increases.

More specifically, we start with two estimators for a hospital's volume per admission: the mean of the admissions in the 5-percent sample and the prediction from a regression of volume per admission on hospital-level and area-level characteristics (e.g., teaching status). These estimators are analogous to the early-season average and the previous season's average. We create a shrinkage estimator in which the weight for the 5-percent mean increases with its precision.

Before laying out the estimator, we must define our notation.[6] Let $Y_{j,t}$ be the average volume per admission for medical staff j (j=1,...,N) in year t (t=1,2), where the casemix-adjusted RVUs for a given admission are weighted by the admission's DRG weight. That is,

$$Y_{j,t} = [\Sigma_i (R_{i,j,t}/w_{i,j,t}) w_{i,j,t}] / \Sigma_i w_{i,j,t} = \Sigma_i R_{i,j,t} / \Sigma_i w_{i,j,t}$$

where $R_{i,j,t}$ is the RVUs of the ith admission and $w_{i,j,t}$ is its relative weight.
(The subscript t is dropped henceforth.)

We specify an error components model in the following regression equation:

$$Y_j = \beta X_j + u_j + e_j$$

where X_j is a vector of explanatory variables and β is a vector of coefficients. We assume that the error terms are independent and

$$u_j \sim N(0, z^2)$$

$$e_j \sim N(0, s^2_j)$$

z^2 is the hospital-level variance (cross-sectionally, not temporally), that is, the variance if each hospital had an infinite number of admissions in a given year. s^2_j represents the component of variance in the jth hospital due to admission-level variance. (The error term associated with s^2_j will be heteroscedastic.)

[6]The notation follows that of Miller and Welch (1993, Appendix A) except that volume per admission is represented by Y instead of by X.

Let

$$y_j = \beta X_j + \phi_j \;;\; \phi_j = u_j + e_j$$

$$\hat{\phi}^2_j = (y_j - \hat{\beta} X_j)^2$$

where the hat symbol indicates an estimate. Given the assumption of independence between the two error terms,

$$E(\phi^2_j) = z^2 + s^2_j$$

$$z^2 = E(\phi^2_j) - s^2_j$$

A reasonable estimator of z^2 is the average deviation over all hospitals, or

$$\hat{z}^2 = [\Sigma_j \hat{\phi}^2_j/N] - [\Sigma_j \hat{s}^2_j/N]$$

where $\Sigma_j \phi^2_j/N$ is approximately the mean square error and s_j is the standard error of Y_j, and[7]

$$\hat{s}^2_j = \frac{\Sigma_j w_{ij}(R_{ij}/w_{ij} - Y_j)^2}{(n_j - 1)\,\Sigma_j w_{ij}}$$

We propose a shrinkage estimate of the volume per admission for hospital j:

$$c_j * (\textit{observed value}) + (1 - c_j) * (\textit{predicted value})$$

$$= (c_j * Y_j) + ((1 - c_j) * \hat{\beta} X_j)$$

where the credibility factor is

$$c_j = \left[\frac{\hat{z}^2}{\hat{z}^2 + \hat{s}^2_j}\right]$$

The greater the standard error (s) of the mean estimator, the less its credibility factor (c). If s^2 is large relative to z^2, c will be close to zero and the shrinkage estimator will largely reflect the regression prediction. On the other hand, if s^2 is small relative to z^2, c will be close to 1.0 and

[7]Whereas the denominator is N, the denominator of the regression-generated mean square error is N-K, where K is the number of independent variables. Because our regression involves several thousand hospitals, this distinction is minor.

Although our software precludes it outside of the equation mode, both ϕ and s should have "hats" in this sentence.

the shrinkage estimator will largely reflect the hospital-specific value. s^2 decreases as the number of admissions for a hospital increases.

The shrinkage estimator has the advantages of less variance than a mean estimator based on a 5-percent sample but more recognition of idiosyncratic factors at the hospital level than a regression prediction. A priori, it is unclear whether the shrinkage estimator over- or underestimates the variance across hospitals based on 100 percent of Medicare admissions in a given year. Below we investigate that empirically.

Regression Specification

Our shrinkage estimator requires explanatory variables. In considering possible explanatory variables, one should keep in mind that we are not testing hypotheses; we are merely trying to predict volume per admission. Therefore, excluding possible explanatory variables that may bias coefficients of other variables is not a concern here. In the interest of simplicity, we specify a regression essentially limited to traditional cost function variables. Note that the dependent variable is in natural form, that is, not logarithmic form.

With one exception discussed below, all of the explanatory variables were defined as they were in Miller and Welch (1993). Three hospital variables were defined in terms of logs, including the bed size, interns-and-residents-to-beds ratio as a measure of teaching activity, and disproportionate share ratio. Four hospital variables were dummy variables: two variables represented rural referral centers and sole community hospitals, respectively. Two other dummy variables pertained to hospital control, one for not-for-profit hospitals and one for state government hospitals. (The excluded category for control was proprietary hospitals. The sole federal government hospital that remained in the database after other screens was dropped.) There were four medical staff variables: physicians per bed, and the percentages of the medical staff that were surgeons, specialists, and RAPs (radiologists, anesthesiologists, and pathologists).

One of the better predictors of a given hospital's volume per admission is the volume per admission of other hospitals in the Metropolitan Statistical Area (MSA), because of the importance of community practice styles (Welch, et al., 1993).[8] Hence, we used MSA-level volume as an explanatory variable.[9] As with the PPS wage index, each state's rural area is treated as if it constituted an MSA. A complication with this variable is that some small MSAs have only one or two hospitals. Then we would be essentially using a hospital's own volume per admission to predict its volume per admission. To avoid this problem, all free-standing MSAs with populations less than 250,000 were combined with rural areas, and all primary MSAs with fewer than 3 hospitals were combined with the largest component of their Consolidated MSA.[10]

---

[8]Because causality will not be inferred, we are not concerned with possible simultaneity, that is, where the dependent variable might influence an independent variable, which in turn influences the dependent variable.

[9]We used the MSA definitions based on the 1990 Census. These definitions were promulgated June 30, 1993.

[10]First metropolitan areas are defined. Then metropolitan areas with populations exceeding 1 million may be subdivided into what is termed "primary MSAs." Then the larger entity is termed a "Consolidated MSA." All metropolitan areas that are not subdivided are termed "free-standing MSAs."

The number of observations (MSAs or rural areas) went from 351 to 208, and there was an average of 24 hospitals per area. The standard error for area-level volume averaged only 1 percent of the national mean volume. Hence, area-level volume has been estimated with precision.[11]

Database

To simulate illustrative policy options, we need a pooled cross-sectional sample of hospitals. Miller and Welch (1993) describes the database construction for 1989. Using the same methodology, we constructed a database for 1991 from the National Claims History (NCH) System. Both files are based on 5-percent samples and have volume per admission by hospital.

As noted above, 1992 RVUs were attached to each year's physician claims, and the 1989 DRG Group was used to develop casemix weights for each year. Because PPS recognizes outlier cases, a medical staff policy might also recognize them. To allow for simulation of policies with (as well as without) outlier provisions, outlier cases were truncated at 2.5 times the DRG mean, and RVUs per admission were calculated both with truncated and with untruncated RVUs. In addition, there are at least two ways to aggregate RVUs per admissions from the hospital-level to, say, the state-level. As explained in Appendix C, we weighted by the number of a hospital's admissions.

We retained only hospitals for which we had complete data: volume in 1989, volume in 1991, and the explanatory variables specified above. Of the 5,076 hospitals used in Miller and Welch (1993), 83 hospitals did not appear in the 1991 HCFA files, perhaps because they closed or merged with another hospital. (These hospitals had only one-third of one percentage point of the admissions in 1989.) Another 37 hospitals were dropped because of very low volume in 1991. (Many of these were Indian Health Service hospitals.) Finally, the sole federal hospital was dropped for simplicity (as mentioned) and one hospital was dropped because of missing data on the medical staff variables. Our final database had 4,954 hospitals. (In addition, we had data for the 12 hospitals in Rhode Island for 1991 but not 1989. These hospitals were not used in some analyses.)

Regression Results

Table 3 presents the regression results. Because the regression's purpose is prediction, not causal inference, the R-square is important and the regression coefficients are less so. This specification explains about two-thirds of the variance, depending on the year.[12]

As discussed above, the shrinkage estimator is a weighted average of the observed and predicted value. Key to calculating the shrinkage estimator is the credibility factor, which has two

[11]The specification here differs from that in Miller and Welch (1993) only in that casemix has been moved from the right- to the left-hand side of the equation, the dependent variable is in natural form, and the area-level volume has replaced the geographic binary variables representing metropolitan area size (i.e., urban-rural location) and region of the country.

[12]We dropped from the regression any hospital with fewer than 10 admissions in our database (about 400 hospitals). These hospitals were retained in the simulation analysis.

components. One component (z^2)--hospital-level variance if there was no admission-level variance--took the value of 19.49 for 1989 and 18.39 for 1991. The other component (s^2_j) was the square of the standard error of the mean estimate for the jth hospital and as such varies by hospital.

The credibility factor ranged from zero to .97. (We set c equal to zero whenever a hospital had fewer than 10 admissions in our database. Then the estimated volume was simply the regression predictor.) The average credibility factor across 4,954 hospitals (unweighted by admissions) was .70. That is, for the typical hospital, the observed value constituted about two-thirds of the shrinkage estimate and the predicted value, about one-third.

Shrinkage Estimates vs. Estimates Based on 100-Percent Data

We compare two sets of estimates of volume per admission for the same hospitals (N=714). One set of estimates is based on 100 percent of the admissions in the seven states discussed in the previous section. We label these "100-percent estimates." The other set is the shrinkage estimates, which are based on a 5-percent national sample. Keep in mind that we wish to simulate a policy based on 100-percent of the admissions. Therefore, we are comparing the 5-percent distribution to the "true" 100-percent distribution. To that purpose, we compare the estimates first in terms of cross-sectional variance and then in terms of temporal variance (change between 1989 and 1991).

Cross-sectionally, the two estimators yield a very similar distribution of volume per admission at the hospital level, as shown in Table 4. The shrinkage estimator using the 5-percent file yields a distribution with slightly smaller tails than the 100-percent estimator. For instance, in 1989 the shrinkage estimate had only .1 percent of hospitals above 1.3 times the mean, whereas the 100-percent estimator had 1.5 percent.[13] The shrinkage estimator appears to slightly underestimate the cross-sectional variance at the hospital level. For a high-cost medical staff policy, the shrinkage estimator would slightly underestimate the savings from lowering payments to staffs in the upper tail of the distribution.

Our comparison of the temporal variance involves the autocorrelation of esiimates in 1989 and 1991. The shrinkage estimates in 1989 and 1991 have a correlation of .880 for these 714 hospitals. (For the 4,954 hospitals in the national database used below to simulate policies, this correlation is .874--essentially the same.) As noted above, the 100-percent estimates in 1989 and 1991 have a correlation of .916 for these 714 hospitals. That is, the unexplained (temporal) variance as percentage of total variance falls from .120 to .084 percent when the shrinkage estimates are replaced by 100-percent estimates--a drop of 30 percent.[14] Whereas the shrinkage estimates underestimate the cross-sectional variance, they overestimate the temporal variance.

[13]The distributions have more hospitals below the mean than above, because the mean is weighted by admissions, whereas an observation in a distribution is not.

[14]The correlations in this paragraph pertain to untruncated values, whereas Table 4 pertains to truncated values. Simulations below use both types of values.

This finding--that there is less year-to-year variance at the hospital level than our shrinkage estimates suggest--has two implications. First, if the 100-percent estimates replaced the shrinkage estimates, we would expect less temporal variance. Second, budget savings from a high-cost medical staff policy might be underestimated by the shrinkage estimator.

In conclusion, the shrinkage estimator

- slightly underestimates the cross-sectional variance and
- somewhat overestimates the temporal variance.

Overall, the similarity of the two distributions in 1989 and 1991 is consistent with the notion that the shrinkage estimator can be used to simulate the impact of policy options. That is, the shrinkage estimator can be used to analyze the impact on hospitals nationally, whereas the 100-percent estimates are available for only 7 states.

B. Illustrative Simulations

This section uses the database developed above to simulate an example of per-case prospective payment to medical staffs and a high-cost medical staff policy.

The examples have two policy parameters in common: First, outliers would be recognized by the policies. Under prospective payment, medical staffs would be paid an additional amount for all physician services beyond an outlier threshold. Under a high-cost policy, physician services beyond the threshold would be excluded from the calculation of each medical staff's performance in any year, measured in terms of casemix-adjusted RVUs per admission. Second, no adjustments (e.g., for teaching status) would be made for hospital type. (See Miller and Welch [1993] for a discussion of outliers. Note that the Clinton Administration's proposal, which is simulated in Welch and Miller (1994) does not recognize outliers but does adjust for teaching and disproportionate share status.

Per-Case Prospective Payment

The prospective payment policy simulated here is simple: The payment rate for each medical staff would reflect the national mean casemix-adjusted RVUs per admission. More precisely, the total payment to a medical staff would be the product of the number of its admissions, its casemix index, its area's geographic adjustment factor in the MFS, and the national casemix-adjusted RVUs per admission. By construction, medical staffs with volume per admission at the national mean would be paid the same under a prospective payment policy as their members would have been paid in aggregate under the status quo. A medical staff with volume per admission of, say, 125 percent of the national mean would be paid 20 percent less under prospective payment than under the status quo. (The inverse of 1.25 is .80.) Conversely, a staff at 75 percent of the national mean would be paid 33 percent more under this policy.

Table 5 shows the number of medical staffs that would gain revenue ("winners") and lose revenue ("losers"). Thirty-one percent of staffs would lose and 69 percent would gain revenue. The winners outnumber the losers, because the payment rate would be based on the admission-weighted national mean and large hospitals (i.e., those with many admissions) tend to have high volume per admission. Almost half of the hospitals are small (i.e., with fewer than 100 beds)

and these hospitals tend to gain. (The number of hospitals and admissions on which these estimates are based is presented below in Table 6.)

Prospective payment would involve considerable redistribution among staffs. Fifteen percent of staffs would lose at least 10 percent of Medicare revenue, whereas 56 percent would gain at least 10 percent of revenue. Going further out on the tails of the distribution, 2 percent of staffs would lose at least 25 percent, whereas 42 percent would gain at least 25 percent.[15]

The mean volume per admission serves as a summary measure of the impact by hospital type. As a group, staffs of government hospitals would gain but staffs of private nonprofit and for-profit hospitals would lose. Staffs of small hospitals would gain at the expense of large ones. Staffs in rural areas would gain at the expense of staffs in large urban areas. Nonteaching and major teaching hospitals would gain at the expense of staffs in minor teaching hospitals. Staffs of disproportionate share hospitals would be hurt slightly relative to other staffs. Staffs in the Northeast states would be penalized the most, whereas staffs in the North Central states would be penalized the least.

Of course, paying staffs at the national mean would not yield budget savings. Alternatively, prospective payment might be made at 95 percent of the national mean, yielding 5 percent savings. Then the above figures can be modified in a straightforward manner. For instance, instead of 31 percent of staffs losing at least some revenue, at least 31 percent would lose at least 5 percent of their revenue.

If prospective payment levels were held constant for several years (except for increases in the conversion factor), budget savings would be a function of how rapidly expenditures would have otherwise increased.[16] Appendix D discusses the growth in volume per admission and the casemix index.

High-Cost Medical Staff Policy

The second policy simulated here is more complex and was described in Welch and Miller (1992). To understand the mechanisms of such a policy, consider a policy implemented in 1995. In 1994, HCFA would determine which medical staffs would be designated "high cost", based on 1993 performance, which is measured in terms of casemix-adjusted RVUs per admission. In 1995, HCFA would apply a withhold for services delivered in 1995 by these high-cost staffs. In 1996, HCFA would return the appropriate portion of the 1995 withhold--penalizing staffs only if

[15]Although Table 5 implicitly presents the distribution of staffs by volume per admission, the intervals require some elaboration. For instance, a staff with a volume per admission of 80 percent of the national mean would gain 25 percent of revenue. That is, its gain would be 20 percent of the national mean or 25 of its own mean. Arithmetically, the inverse of 80 percent (4/5) is 125 percent, which exceeds 1.0 by 25 percent.

To take another example, a staff with a volume per admission of 133 percent of the national mean would lose 25 percent of revenue. Arithmetically, the inverse of 4/3 is 75 percent.

[16]Similarly, if a high-cost medical staff policy had a limit (in terms of RVUs per admission) that did not change for several years, budget savings would be a function of how rapidly RVUs per admission would otherwise have increased.

they were designated high cost on the basis of 1993 performance and were high cost again based on performance in the withhold year (1995).[17]

For the purposes of this simulation, we must specify several parameters. High-cost staffs would be defined as those with volume per admission at or above 120 percent of the median (or 104 percent of the weighted mean.)[18] The withhold would be 10 percent of payment; thus, the maximum any staff would be penalized is 10 percent. As noted above, an outlier policy is presumed but no adjustments would be made for hospital type.

Table 6 presents estimates of the impact by hospital type of such a policy implemented in 1991. The 1989 data indicate that 23 percent of medical staffs would be designated as "high cost" and would have a withhold applied to their payments in 1991. The 1991 data indicate that 17 percent of all staffs would be penalized with some loss of payment. About a quarter of high-cost staffs would have the withhold returned in full in 1992 (i.e., 1 - (16.8/23.3)). Across all high-cost staffs, sixty percent of the withhold amounts would be retained by Medicare (not shown).

In contrast to a prospective payment policy, there are necessarily no winners. However, there are necessarily some budget savings. The same hospital types would lose under this policy as under prospective payment. However, no staff would lose more than 10 percent under the illustrative high-cost policy, whereas about 100 (or 2 percent) would lose more than 25 percent under the illustrative prospective payment.

When policy options are under serious consideration, many variants are usually simulated. This database was constructed to provide maximum flexibility in simulating various options. Under prospective payment, it could simulate different payment levels. Under a high-cost policy, it could simulate different high-cost limits and different withhold amounts. Under both policy options, it could simulate policies with and without outliers and policies that adjust for teaching status, disproportionate share status, and other characteristics. The two simulations presented here are only illustrative of the uses of this database.

As part of its Health Security Act, the Clinton Administration has proposed a high-cost medical staff policy. Welch and Miller (1994) simulate the impact of this proposal.

[17]The RVUs for a given HCPCS code may change from year to year. Administration of a policy would be easier if, for instance, 1994 RVUs were used to measure performance in 1994 and 1996 RVUs were used to measure performance in 1996. (We used 1992 RVUs for both 1989 and 1991 data.) Across-the-board changes in RVUs, as occurred between 1992 and 1993, might necessitate an offsetting modification of the limit.

[18]The volume per admission for the median hospital was 31.6 in 1989 and 32.0 in 1991. The median is unweighted by admissions and has been calculated from the shrinkage estimates.

REFERENCES

Cahill, K.R., J. Reuter, K. Langwell, and R. Shin (1989) "Prospective Budget for Medicare's Physician Service," CRS: Washington, DC.

CHER [Center for Health Economics Research] (1991) "Summary of 100 Percent Part B Data Availability from Health Economics Research, Inc.," May.

Duan, N., W. G. Manning Jr., and J. P. Newhouse (1989) "Credibility Method for Setting Insurance Premiums," manuscript, May.

Efron, B. and C. Morris (1977) "Stein's Paradox in Statistics," Scientific American 236 (May): 119.

Miller, M. E., S. L. Hui, W. M. Tierney, and C. J. McDonald (1993) "Estimating Physician Costliness: An Empirical Bayes Approach," Medical Care 31 (May, Supplement): 16-28.

Miller, M. E., and W. P. Welch (1993) "Medical Staff Risk Pools: Technical Report," Urban Institute working paper #6210-01, March.

[PPRC] Physician Payment Review Commission (1990) PPRC, Annual Report to Congress. PPRC: Washington, DC.

[ProPAC] Prospective Payment Assessment Commission (1993) Report and Recommendations to the Congress, March 1, 1993. ProPAC: Washington, DC.

Welch, W. P., and M. E. Miller (1992) "High-Cost Medical Staffs: A Policy Option for Controlling the Volume of Physician Services in the Hospital," Urban Institute working paper #6210-04, September.

- - (1993) "Toward Profiling Physicians for Inpatient Services: Florida," Urban Institute working paper #6210-02, July.

- - (1994) "The High-Cost Hospital Medical Staff Proposal in the Health Security Act: Distributional Impacts," Urban Institute working paper #6210-08, May.

Welch, W. P., M. E. Miller, H. G. Welch, et al. (1993) "Geographic Variation in Expenditures for Physicians' Services in the United States," NEJM 328 (March 4): 621-627.

Table 1

Regression of 1989 Volume per Admission on 1988 Volume per Admission

Description of Definition	Intercept	Slope	R^2	r
Stay Only, Untruncated	1.4 (.79)	.978 (.0091)	.94	.97
Stay Only, Truncated	1.6 (.82)	.977 (.0094)	.94	.97
Includes Windows, Untruncated	2.7 (.82)	.967 (.0093)	.94	.97
Includes Windows, Truncated	2.8 (.82)	.967 (.0093)	.94	.97

N = 733
Standard errors are in parentheses.
Volume per admission is casemix-adjusted RVUs per admission. Each volume-per-admission variable was normalized to 100 by its year-specific grand mean (weighted by admissions). The regression, however, is not weighted.

Table 2

Distribution of First Differences in Volume per Admission

(Normalized Variables, 1989 minus 1988)

Hospitals	Standard Deviation	Percentiles of the Distribution							N
		5%	10%	25%	50%	75%	90%	95%	
All	6.3	-10.1	-7.2	-3.4	-0.1	3.3	6.1	8.3	733
In Hospitals with >50 Beds	6.0	-9.0	-7.0	-3.2	0.2	3.7	6.1	8.2	545
>110% of the National Mean in 1988	6.8	-12.8	-10.0	-6.1	-1.7	2.8	5.5	8.2	120

Volume per admission pertains to the stay, and outlier cases have not been truncated. Each volume-per-admission variable was normalized to 100 by its year-specific grand mean.

Table 3
Regression of Volume per Admission
for Shrinkage Estimates

Independent Variables	1989	1991
Bed Size (log)	4.02**	4.00**
	(31.58)	(31.75)
Medicare-Defined Characteristics		
Teaching Status (log)	-18.63**	-17.36**
	(17.30)	(16.76)
Disproportionate Share Status (log)	-3.54**	-0.67
	(2.95)	(0.57)
Rural Referral Center	1.57**	1.43**
	(3.91)	(3.69)
Sole Community Hospital	-0.04	-0.21
	(0.11)	(0.62)
Staff Composition		
Staff/Bed Ratio (log)	1.95**	1.97**
	(14.31)	(14.97)
Specialists/Staff Ratio	1.45*	1.36*
	(2.24)	(2.16)
Surgeons/Staff	7.01**	5.14**
	(7.93)	(6.10)
RAPs/Staff	1.61	-0.68
	(1.60)	(0.71)
Ownership		
Proprietary	0.84**	0.68**
	(3.62)	(3.04)
State Government	-0.78**	-0.80**
	(3.88)	(4.05)
Area-Level Volume per Admission	0.80**	0.82**
	(40.42)	(42.76)
Intercept	-15.34	-14.96
R^2	.64	.68
F	678	817
Mean Square Error	26.79	24.94
N	4540	4515

** Significant at 99 percent confidence level.
* Significant at 95 percent confidence level.
Absolute t-value in parentheses.
The dependent variable is casemix-adjusted RVUs per admission and pertains to stay only. Outliers have been truncated. Disproportionate share status is represented by one plus the disproportionate share percentage and teaching status is represented by one plus the hospital's intern-and-resident-to-bed-ratio. The category representing non-government not-for-profit hospitals is the excluded group under ownership.

rdm g:\welchp\6210\stab\shrkreg.t3 4/25/94 10:51am

Table 4
Distribution of Hospitals
by Volume per Admission
in Two Files, 1989 and 1991

Range of Normalized Variable	1989		1991	
	100% File	5% File	100% File	5% File
<.7	30.1	29.1	33.8	32.9
.7-.8	12.7	11.2	10.8	11.8
.8-.9	11.9	12.7	11.9	12.7
.9-1.0	10.8	10.8	16.1	14.7
1.0-1.1	14.4	10.8	11.2	13.0
1.1-1.2	10.8	10.8	7.6	7.3
1.2-1.3	3.8	2.7	5.0	4.9
1.3<	1.5	0.1	3.6	2.7
Mean				
Unnormalized	35.1	37.5	38.9	38.8
Normalized	1.00	1.00	1.00	1.00
Number of Hospitals	714	714	714	714

Casemix-adjusted RVUs per admission pertains to services delivered during the stay, and values exceeding the outlier threshold were truncated. This variable was normalized separately for each data set and each year for the same 714 hospitals. The means used to normalize were weighted by the number of admissions. The 5-percent file constitutes shrinkage estimates.

Table 5
Impact of an Illustrative
Per-Case Prospective Payment
by Hospital Type[1]

Hospital Type	Volume per Admission[2]	Percentage of Staffs in 1991					
		Gain			Lose		
		+25%	10-25%	0-10%	0-10%	10-25%	+25%
United States	**1.00**	**41.5%**	**14.7%**	**13.2%**	**16.0%**	**12.9%**	**1.8%**
Control							
Private Nonprofit	1.02	30.8	15.9	15.6	19.3	16.3	2.2
Government	0.85	73.7	11.3	7.7	5.5	1.9	0.1
For-Profit	1.03	24.7	15.8	14.1	22.5	19.5	3.3
Bed Size							
<50	0.64	92.3	5.0	1.7	0.8	0.2	0.0
50-100	0.80	62.6	18.4	10.2	5.6	3.1	0.1
101-200	0.96	19.1	26.3	18.1	19.8	15.0	1.7
201-300	1.06	3.6	12.2	19.6	32.2	26.6	5.9
301-500	1.07	3.9	8.1	19.7	33.2	31.1	3.9
>500	1.05	3.6	9.9	26.6	32.4	26.1	1.4
Urban/Rural Location							
Rural	0.80	75.0	15.5	6.0	2.6	0.9	0.0
Other Urban	1.02	16.7	18.1	23.0	26.0	14.4	1.9
Large Urban	1.10	6.0	8.8	15.7	30.2	34.3	5.0
Teaching							
Nonteaching	0.96	50.9	15.2	11.2	12.2	9.3	1.3
Minor Teaching	1.07	6.6	11.9	18.6	31.2	27.7	4.0
Major Teaching	0.99	17.0	17.0	28.6	21.8	14.1	1.5
Disproportionate Share Status							
No	0.99	48.9	14.8	11.4	12.9	10.7	1.4
Yes	1.02	21.2	14.4	18.4	24.5	18.9	2.7
Region							
New England	0.95	23.5	27.9	25.5	20.1	2.9	0.0
Middle Atlantic	1.14	6.1	10.0	14.5	19.8	38.3	11.2
South Atlantic	1.02	30.4	15.9	14.0	17.2	19.1	3.4
East North Central	1.00	35.3	17.7	14.1	17.5	15.5	0.0
East South Central	0.92	55.7	14.4	12.8	14.8	2.3	0.0
West North Central	0.88	73.1	10.2	7.8	7.5	1.3	0.0
West South Central	0.97	49.6	12.4	11.9	16.1	9.5	0.6
Mountain	0.91	64.4	11.3	10.7	9.4	4.2	0.0
Pacific	0.99	31.8	17.1	15.1	22.7	12.7	0.7
Puerto Rico	0.85	40.0	37.5	15.0	7.5	0.0	0.0

[1] This policy option would pay each medical staff at the national mean. Values exceeding the outlier threshold have been truncated. Casemix was adjusted for, but teaching and disproportionate share status were not.

[2] Casemix-adjusted RVUs per admission in 1991 normalized to the national mean of 37.1.

Table 6
Impact of an Illustrative
High-Cost Medical Staff Policy
by Hospital Type[1]

Hospital Type	Volume per Admission[2]	% of All Medical Staffs: Staffs High-Cost in 1989	% of All Medical Staffs: Staffs Penalized in 1991	Decrease in Payments to All Staffs (%)	Database: Number of Hospitals	Database: Number of Admissions 1991
United States	**1.00**	**23.3%**	**16.8%**	**3.1%**	**4,954**	**458,571**
Control						
Private Nonprofit	1.02	27.9	20.5	3.1	2,933	342,253
Government	0.85	5.8	2.9	1.1	1,333	65,687
For-Profit	1.03	38.1	28.5	4.4	688	50,631
Bed Size						
<50	0.64	0.5	0.3	0.1	1,135	18,533
50-100	0.80	6.3	3.1	0.6	1,172	45,956
101-200	0.96	29.5	19.4	2.1	1,137	93,652
201-300	1.06	47.9	36.6	3.4	674	100,951
301-500	1.07	52.1	41.5	3.7	614	128,110
>500	1.05	44.1	33.7	3.4	222	71,369
Urban/Rural Location						
Rural	0.80	3.7	1.1	0.4	2,329	102,606
Other Urban	1.02	34.2	22.5	2.8	1,422	191,420
Large Urban	1.10	48.5	40.8	4.3	1,203	164,545
Teaching						
Nonteaching	0.96	18.2	12.7	3.0	3,850	257,393
Minor Teaching	1.07	46.4	35.4	3.6	898	160,724
Major Teaching	0.99	18.4	13.1	1.6	206	40,454
Disproportionate Share Status						
No	0.99	20.0	14.4	3.3	3,635	293,634
Yes	1.02	32.6	23.7	2.7	1,319	164,937
Region						
New England	0.95	23.5	12.7	0.7	204	24,018
Middle Atlantic	1.14	44.2	42.4	5.1	509	76,316
South Atlantic	1.02	31.1	25.7	4.0	756	85,747
East North Central	1.00	19.2	14.1	2.5	782	82,463
East South Central	0.92	17.1	7.7	1.7	438	38,742
West North Central	0.88	7.2	3.8	1.3	677	35,704
West South Central	0.97	26.7	16.0	3.4	688	49,898
Mountain	0.91	20.4	8.4	2.4	309	18,404
Pacific	0.99	23.0	17.0	2.1	551	44,124
Puerto Rico	0.85	0.0	0.0	0.0	40	3,155

[1] This policy option would define medical staffs with volume per admission above 120% of the national median to be high cost. Ten percent of payments to members of these staffs would be withheld. Values exceeding the outlier threshold were truncated. Casemix was adjusted for, but teaching and disproportionate share status were not.

[2] Casemix-adjusted RVUs per admission normalized to the national mean of 37.1.

Appendix A: Levels Vs. Growth Rates

The issue of instability has been raised in the context of state-level VPSs. Cahill, et al. (1989) reported a correlation of .01 between volume and intensity growth in one year and growth in another year. Similarly, PPRC (1990) found "large year-to-year changes in reported total Part B payments [that] follow a clear 'seesaw' pattern, with high-growth years followed by low-growth years and vice versa." In contrast, this memo finds a correlation of .97 for hospital-level volume and intensity levels. We believe this large difference in stability represents the difference between growth rates and levels, not between state-level and hospital-level measures.

Consider a generic measure of volume and intensity, X. Suppose X has no secular trend, either at the national level (perhaps it has already been normalized) or at the unit level (e.g., state or hospital level). Volume in year t is X plus a year-specific error term, e_t, which we assume has an autocorrelation of zero. The volume levels in years 0 and 1--$(X + e_0)$ and $(X + e_1)$--are positively correlated. In fact, if the cross-sectional variance of X across states (or hospitals) is large relative to the variance of e_t, the correlation between levels will be large and can plausibly approach 1.0.

The volume growth rates in years 1 and 2 are

$$[(X + e_1)/(X + e_0)] - 1 \quad \text{and} \quad [(X + e_2)/(X + e_1)] - 1 .$$

(Multiplied by 100, these yield percentage growth rates.) When e_1 is positive, the growth rate in year 1 is likely to be positive and the growth rate in year 2 is likely to be negative. In fact, in this simple model of no secular trend at the state (or hospital) level, the growth rate in one year is negatively correlated with the growth rate in the previous year.[1] A substantial secular trend by state (i.e., positive autocorrelation) is necessary to overwhelm this statistical property.

Under the simplifying assumption of no secular trend, volume and intensity levels are positively correlated but volume and intensity growth rates are negatively correlated. Hence, it is unsurprising that Cahill, et al. (1989) and PPRC (1990) found that volume and intensity growth rates are unstable, whereas we find that volume and intensity levels are stable.

The policy implications are important: instability is a much more serious problem for a staff-level VPS growth standard than for a per-admission budget or a high-cost medical staff policy.

[1] In some preliminary analysis of deflated charges per admission, we found the 1986-87 growth rate and 1987-89 growth rate to be negatively correlated, whereas the 1986, 1987, and 1989 charges-per-admission levels were positively correlated.

Appendix B: Random Fluctuation at the Admission Level

We need a summary measure of the impact of random fluctuations on the year-to-year variation in volume per admission at the hospital level. Although nationwide growth is a motivation for policy consideration, it is changes relative to the grand mean that represents instability. Our 7-state database has volume per admission for each hospital for each of two years (1988 and 1989). This appendix investigates the percentage of year-to-year variation for each hospital that results from random fluctuation at the admission level versus variation that can be considered systematic.

Let $X_{j,t}$ be the volume per admission for the jth hospital in year t (t=1,2), where the casemix-adjusted RVUs for admission is weighted by the admission's DRG weight. That is,

$$X_{j,t} = [\Sigma_i (R_{i,j,t}/w_{i,j,t})w_{i,j,t}]/\Sigma_i w_{i,j,t} = \Sigma_i R_{i,j,t}/\Sigma_i w_{i,j,t}$$

where $R_{i,j,t}$ is the RVU of the ith admission and $w_{i,j,t}$ is the relative weight of its DRG. Let $s_{j,t}$ be the standard error of that average, where

$$s^2_{j,t} = [\Sigma_i w_{i,j,t} ((R_{i,j,t}/w_{i,j,t}) - X_{j,t})^2]/[(n_{j,t}-1)\Sigma_i w_{i,j,t}]$$

Expansion yields

$$s^2_{j,t} = [\Sigma_i (R^2_{i,j,t}/w_{i,j,t}) - (X^2_{j,t} \Sigma_i w_{i,j,t})]/[(n_{j,t}-1)\Sigma_i w_{i,j,t})]$$

The hospital-level change in volume per admission is

$$\Delta_j = X_{j,2} - X_{j,1}$$

($X_{j,2}$ has already been adjusted for any nationwide change from the previous year. This involved dividing volume per admission in 1989 by 1.025.) The variation in Δ_j that is attributable to randomness is

$$s^2_{j,1} + s^2_{j,2}$$

(By construction, the samples in year 1 and year 2 are independent.) The standard error of Δ_j is

$$[s^2_{j,1} + s^2_{j,2}]^{1/2}$$

The average standard error across hospitals is .97 RVUs per admission. Dividing by the mean for 1988 (35.65 RVUs per admission) yields 2.7 percentage points.

In contrast, the standard deviation of Δ_j is 2.2 RVUs per admission, or 6.3 percentage points. Hence, about 18 percent ($(2.7/6.3)^2$) of the within-hospital between-year variance in volume per admission can be attributed to admission-level fluctuation. The remaining variance is systematic.

Appendix C: Aggregation of Data on Volume per Admission

Summary: We aggregate volume per admission following two rules:

- When aggregating admission-level volume to the hospital level, we divide the sum of the RVUs by the sum of the weights.
- When aggregating volume per admission from the hospital level to any geographic level (i.e., MSA, state, or nation) or to the hospital-type level (e.g., teaching hospitals), we take the average weighted by each hospital's admissions.

Aggregation to the hospital level: Let R_{ij} be the RVUs of the ith admission in the jth hospital and w_{ij} be the relative weight of that admission. The volume per admission of the jth hospital is calculated as

$$Y_j = \frac{\Sigma_i R_{ij}}{\Sigma_i w_{ij}} \quad (1)$$

Welch and Miller (1993, Appendix A) explained the rationale for this aggregation approach and showed how to test for significance with it. (The present appendix builds on that discussion and follows its notation.)

Aggregation to higher levels: There are at least two ways to aggregate hospital-level data. One approach calculates the average of hospital-level volume, weighting by the number of admissions:

$$Y_1 = \frac{\Sigma_j Y_j * N_j}{\Sigma_j N_j} \quad (2)$$

where N_j is the number of admissions in our sample to the jth hospital.

A second approach weights by the number of admissions times the casemix index:

$$Y_2 = \frac{\Sigma_j Y_j * (N_j * C_j)}{\Sigma_j (N_j * C_j)} \quad (3)$$

where C_j (the casemix index of the jth hospital) is calculated as $\Sigma_i w_{ij}/N_j$.

The second approach is analogous to our aggregation to the hospital level (eq. 1). To see this, substitute C_j into the equation 3, yielding

$$Y_2 = \frac{\Sigma_j \, Y_j \, (\Sigma_i \, w_{ij})}{\Sigma_j \, (\Sigma_i \, w_{ij})} \qquad (4)$$

Then if Y_j (eq. 1) is substituted into equation 4, $\Sigma_i \, w_{ij}$ cancels out in the numerator, yielding

$$Y_2 = \frac{\Sigma_j \, \Sigma_i \, R_{ij}}{\Sigma_j \, \Sigma_i \, w_{ij}} \qquad (5)$$

Equation 5 is analogous to equation 1, which defines volume per admission at the hospital level.

The second approach (Y_2) is mathematically elegant, in contrast to the first approach (Y_1), which is awkward. If one uses the first approach to aggregate from the hospital level to the state level and then to national level, the mean will not necessarily be the mean resulting from aggregating directly from the hospital level to the national level.

However, weighting by admissions times the casemix index is difficult to articulate, whereas weighting by admissions is common (e.g., ProPAC). Hence, for simplicity when aggregating above the hospital level, we weight by admissions (as in Y_1).[1]

Weighting by the sum of the weights (as in Y_2) yields a slightly higher mean casemix-adjusted RVUs per admission (39.96) than weighting by admissions (39.34) in 1991. (These figures pertain to intensity for stay, untruncated.) This results from large hospitals having higher volume per admission and higher casemix indices (Y_j and C_j). However, both approaches yield growth rates of about 1.9 percent per year.

[1]When aggregating RVUs by type-of-service (TOS) categories above the hospital level, we weight by admissions times casemix (as in Y_2). For certain TOS categories, casemix can vary dramatically across hospitals.

Appendix D: Growth in Volume per Admission and the Casemix Index[1]

Estimate of the growth rate

Although not necessary for distributional simulations, an estimate of the volume-per-admission growth rate could be key for an estimate of budgetary savings from a specific policy option.

To examine this, we calculated the means for two variables in two years:[2]

year	RVUs per admission	casemix index
1989	39.17	1.017
1991	41.10	1.028
average annualized growth rate	2.4%	0.5%

Hence, casemix-adjusted RVUs per admission grew at an annualized rate of 1.9 percent.

Although the physician casemix index increased .5 percent per year between 1989 and 1991, the PPS casemix index increased 2.2 percent between 1989 and 1990 and 2.7 percent the following year (ProPAC, March 1993, p. 70). Our previous work has found PPS relative weights and physician relative weights to be highly correlated. More precisely, a DRG-level regression of the physician weight on the PPS weight and a medical-surgical DRG indicator explained three-quarters of the variance. Hence, one would expect increases in the two casemix indices to be similar.

This issue is important: if our estimate of the increase in casemix is too low, our estimate of the increase in casemix-adjusted RVUs per admission would be too high. Estimates of budget savings from medical staff policies might be too high. However, as laid out below, the difference in rates appears to derive from different index forms. In fact, the PPS "casemix index" might better be termed an "expenditure index."

Laspeyres Price Index

Because of the familiarity of price indices, it is useful to start by writing out a common price index, Laspeyres. If P_{ti} represents the price of good i in year t and Q_{ti} represents its quantity, a Laspeyres price index for year 1 can be written as follows:

---

[1]We are grateful to Stuart Gutterman of the staff of ProPAC for explaining the subtleties of the PPS index to us.

[2]Hospitals were weighted by admissions. The RVUs per admission were not adjusted for casemix, pertain to the stay only, and were not truncated for outliers. Observed values (as distinct from shrinkage estimates), were used. N=4,954 because 1989 data was unavailable for the 12 Rhode Island hospitals.

$$\frac{\Sigma_i P_{1i} Q_{0i}}{\Sigma_i P_{0i} Q_{0i}}$$

Physician Casemix Index

The change in our casemix index (for inpatient physician services) is analogous to a Laspeyres price index. This increase can be written as follows:

$$\frac{\Sigma_i P_{0i} Q_{1i}}{k \Sigma_i Q_{1i}} \text{ divided by } \frac{\Sigma_i P_{0i} Q_{0i}}{k \Sigma_i Q_{0i}}$$

where in this context P_{ti} represents the average cost (or RVUs) of DRG i in year t, Q_{ti} represents the number of admissions in this DRG, and k represents a constant that normalizes the index to 1.0 in the base year. (Relative weights are represented as P_{0i}/k.) If we make the simplifying assumption that there is no change in total admissions (i.e., $\Sigma_i Q_{0i} = \Sigma_i Q_{1i}$), although there may be shifts in admissions among DRGs, the increase in the casemix index simplifies to

$$\frac{\Sigma_i P_{0i} Q_{1i}}{\Sigma_i P_{0i} Q_{0i}}$$

Mechanically, the difference between our index and the Laspeyres index is only in terms of the year of the quantity in the numerator (Q_{1i} in the physician index vs. Q_{0i} in the Laspeyres index). Conceptually, the difference is the following: the Laspeyres price index measures the change in prices, holding the quantities constant; whereas the casemix index measures the change in quantities, holding the prices constant.

PPS Casemix Index

The PPS casemix index is more complex. For exposition, we consider the index concretely by discussing it in terms of 1993 and 1994. Using prices and quantities for 1991, HCFA calculates the relative weights to be used in 1993 for each DRG:

$$\rho_{91i} = \frac{P_{91i}}{\Sigma_i P_{91i} Q_{91i} / \Sigma_i Q_{91i}}$$

(Relative weights are calculated the same way for the physician casemix index: k in the physician casemix index equals the denominator here.)

A year later HCFA calculates the (tentative) weights to be used in 1994:

$$\rho'_{94i} = \frac{P_{92i}}{\Sigma_i P_{92i} Q_{92i} / \Sigma_i Q_{92i}}$$

It also calculates the mean of the weights (ρ_{93i}) that are based on 1991 data but are used in 1993, weighting by the 1992 quantities:

$$M_{92} = \frac{\Sigma_i \rho_{93i} Q_{92i}}{\Sigma_i Q_{92i}}$$

This measures the change in costs due to shifts in quantities between 1991 and 1992. The actual weights used in 1994 are calculated as

$$\rho_{94i} = \rho'_{94i} M_{92}$$

which ensures the following relationship:

$$\frac{\Sigma_i \rho_{94i} Q_{92}}{\Sigma_i Q_{92}} = \frac{\Sigma_i \rho_{93i} Q_{92i}}{\Sigma_i Q_{92i}}$$

That is, the relative weights used in 1993 and those used in 1994 are constructed to have the same mean (weighting by 1992 quantities).

The increase in the casemix index is calculated as

$$\frac{\Sigma_i \rho_{94i} Q_{94i}}{\Sigma_i Q_{94i}} \text{ divided by } \frac{\Sigma_i \rho_{93i} Q_{93i}}{\Sigma_i Q_{93i}}$$

To simplify this as we simplified the change in the physician casemix index, we assume no change in the total admissions, yielding

$$\frac{\Sigma_i \rho_{94i} Q_{94i}}{\Sigma_i \rho_{93i} Q_{93i}}$$

By way of contrast, the increase in the physician casemix index can be written as follows:

$$\frac{\Sigma_i \rho_{93i} Q_{94i}}{\Sigma_i \rho_{93i} Q_{93i}}$$

where year 0 is 1993. Note that $\rho_{0i} = P_{0i}/k$. The PPS index differs from the physician index in that a common set of prices or costs is not used in the numerator and denominator. Conceptually, the change in the PPS casemix index measures changes in both the distribution of

admissions over DRGs and the relative costs of DRGs. The index should be seen as an expenditure index instead of a casemix index.

As noted above, the PPS index increases faster than the physician index. This is because not only are admissions shifting to high cost DRGs (perhaps because of upcoding) but also relative costs of high-cost DRGs are increasing (S. Gutterman, personal communication).

When economists measure price changes, they hold quantities constant. When we measure casemix changes (i.e., quantity shifts), we hold relative prices (i.e., costs or RVUs) constant, whereas the PPS index does not. We believe that our form of a casemix index more closely matches the general notion of such an index.

CMS LIBRARY
3 8095 00012863 3